Prospect Enhancement Tools are (Possibly) Stealing Our Data

Luke Lighthart - Advised by Sid Stamm

Motivation

- Privacy is an important right
- Violations of this right should not be taken lightly
- Discovering and reporting privacy violations
- Prospect enhancement tools (PET) have shown potential to abuse the privacy of anyone in the general population and research into the area has not been conducted to this date.

Background

- Prospecting is an integral part of the sales/recruitment process
- However many of the prospective clients do not directly consent to the collection and processing of their data, or are unaware that their data is being collected
- This collection and processing of data may have other side effects than what the PET companies intend for it (e.g. supporting surveillance capitalism)

Public Data Collection

- Public data is collected from the internet using a web-crawler and manual collection
- Simulates a sales team gathering public data or “hearing” about a potential client
- Used as input to the extensions

Designing Tests

- Each tool is broken down into its respective features and what kinds of data that feature uses
- A test that covers those features’ data types must be developed.

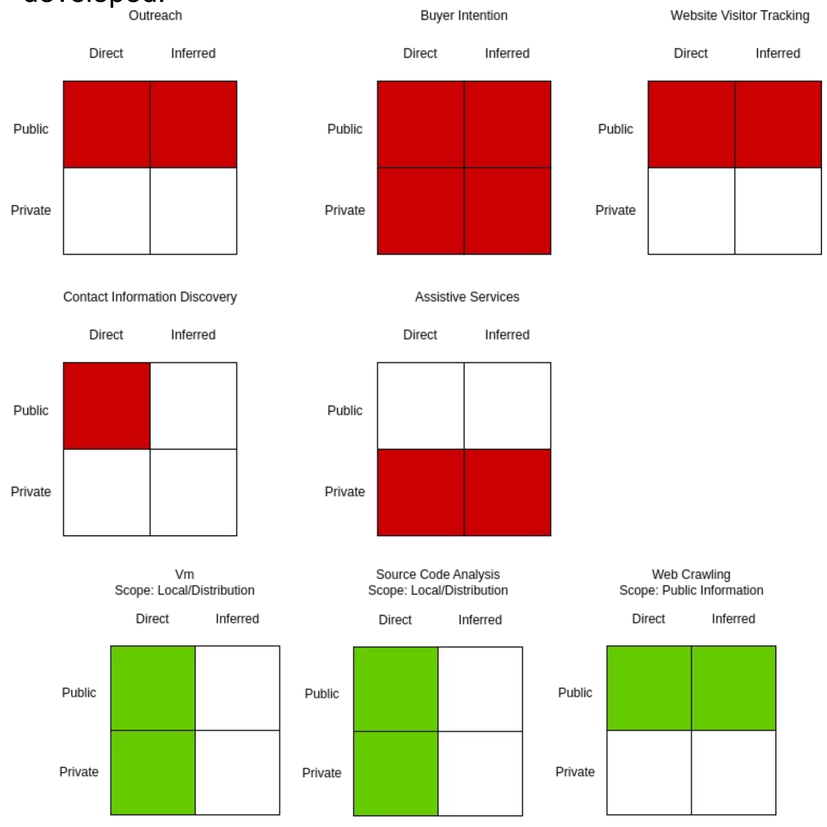

Figure 1: Features, and Tests by Data Type covered

Scoring

- Scoring is based on a tree model (see paper for full model)
- Base scores are assigned to leaf nodes and the final score is the sum of the children

$$S(N) = \alpha_N \left(\sum_{i \in Sub} S(i) \right)$$

Figure 3: Assignment of scores to non-leaf nodes

VM Experiment

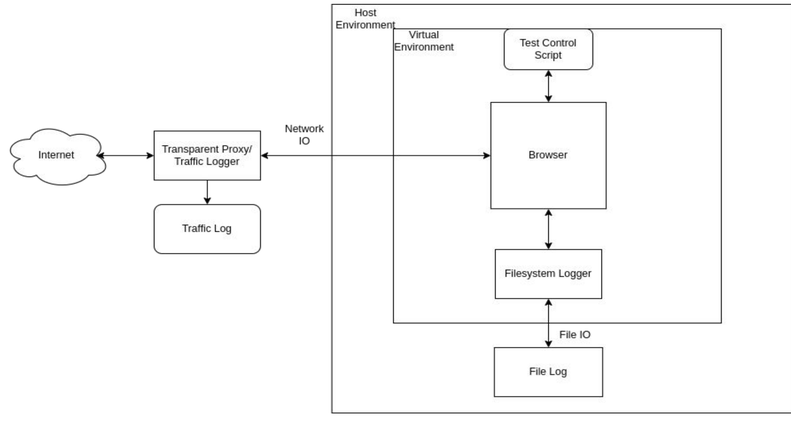

Figure 2: Vm Experiment Architecture

Tools Being Tested

Top 3 that offer browser extensions:

Datanyze

Overloop

Kaspr

Recommendations

- Consent before any kind of data collection
- Easily understandable Privacy Policies
- No distribution of data to third parties without accepting responsibility of what’s happening to it
- Minimal set data collection